AD-A035 387

ON USING SCENARIOS IN THE EVALUATION OF
COMPLEX ALTERNATIVES

DECISIONS AND DESIGNS, INCORPORATED
McLEAN, VIRGINIA

DECEMBER 1976

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR 76-17	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) On Using Scenarios in the Evaluation of Complex Alternatives		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Michael F. O'Connor Ward Edwards		8. CONTRACT OR GRANT NUMBER(s) Prime Contract: N00074-76-C-0074 Subcontract: N00075-030-0711
9. PERFORMING ORGANIZATION NAME AND ADDRESS Decisions and Designs, Incorporated 8400 Westpark Drive McLean, Virginia 22101		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217		12. REPORT DATE December, 1976
		13. NUMBER OF PAGES 71
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office) Defense Advanced Research Projects Agency (ARPA) 1400 Wilson Boulevard Arlington, Virginia 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Support for this research performed by Social Science Research Institute was provided by the Advanced Research Projects Agency of the Department of Defense and was monitored under Contract N00014-76-C-0074 with the Office of Naval Research, under subcontract from Decisions and Designs, Inc.

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Alternative	Dimension	Tweaking
Attribute	Evaluation	Utility
Choice	Option	Variable
Decision	Probability	
Design	Scenario	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The deployment of large military (or social) systems will occur in an uncertain future, and decisions concerning the design of and choice among alternative systems must therefore involve projections of the expected outcomes of system deployment in such uncertain futures. An attempt to represent the large number of potential alternative futures using an event diagram will result in the bushy-mess problem, often encountered in applications of decision analytic techniques, in which the number of different branches in a probability (or decision) tree is prohibitively large.   (cont.)

An alternative to attempting to exhaustively characterize the future in a system evaluation is to evaluate the alternative systems in a reasonable number of hypothetical scenarios. Such scenarios should satisfy two objectives. They should validly represent the future world in which the systems will be deployed. They should also discriminate between the alternative systems in terms of the utility of system deployment. It need not be the case, of course, that the same set of scenarios can satisfy both objectives. The design and choice of scenarios therefore can be a critical evaluation step.

This report is intended as a general discussion of the problem of the use of such scenarios in system evaluation. Specifically addressed are the issues of the definition and characterization of the scenario problem, the proper selection of scenarios, the evaluation of scenario probabilities, and the use of scenarios in both system design and in choosing among a specified set of alternative systems.
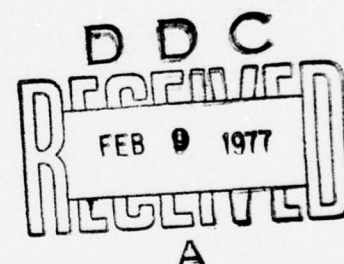
ADA035387

ON USING SCENARIOS IN THE EVALUATION OF COMPLEX ALTERNATIVES

DECISIONS AND DESIGNS INCORPORATED

Michael F. O'Connor
Ward Edwards

D D C

FEB 9 1977

A

ADVANCED ⊕ARPA⊕ DECISION TECHNOLOGY PROGRAM

The objective of the Advanced Decision
Technology Program is to develop and transfer
to users in the Department of Defense advanced
management technologies for decision making.
These technologies are based upon research
in the areas of decision analysis, the behavioral
sciences and interactive computer graphics.
The program is sponsored by the Cybernetics
Technology Office of the Defense
Advanced Research Projects Agency and
technical progress is monitored by the Office
of Naval Research — Engineering Psychology
Programs. Participants in the program are:

Decisions and Designs, Incorporated
The Oregon Research Institute
Perceptronics, Incorporated
Stanford University
The University of Southern California

Inquiries and comments with
regard to this report should be
addressed to:

Dr. Martin A. Tolcott
Director, Engineering Psychology Programs
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

or

LT COL Roy M. Gulick, USMC
Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209

TECHNICAL REPORT 76-17

ON USING SCENARIOS IN
THE EVALUATION OF COMPLEX ALTERNATIVES

Michael F. O'Connor
Decisions and Designs, Incorporated

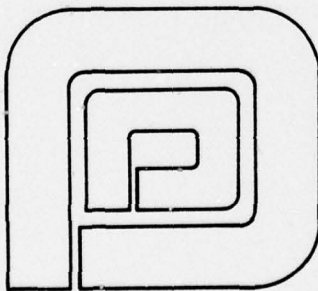and

Ward Edwards
Social Science Research Institute
University of Southern California

December, 1976

Sponsored by

DECISIONS and DESIGNS, INC.

Suite 600, 8400 Westpark Drive
McLean, Virginia 22101
(703) 821-2828

1b

ON USING SCENARIOS IN THE EVALUATION
OF COMPLEX ALTERNATIVES

SUMMARY

The deployment of large military (or social) systems
will occur in an uncertain future, and decisions concerning
the design of and choice among alternative systems must
therefore involve projections of the expected outcomes of
system deployment in such uncertain futures. An attempt to
represent the large number of potential alternative futures
using an event diagram will result in the bushy-mess problem,
often encountered in applications of decision analytic tech-
niques, in which the number of different branches in a
probability (or decision) tree is prohibitively large.

An alternative to attempting to exhaustively characterize
the future in a system evaluation is to evaluate the alter-
native systems in a reasonable number of hypothetical scenarios.
Such scenarios should satisfy two objectives. They should
validly represent the future world in which the systems will
be deployed. They should also discriminate between the
alternative systems in terms of the utility of system deploy-
ment. It need not be the case, of course, that the same set
of scenarios can satisfy both objectives. The design and
choice of scenarios therefore can be a critical evaluation
step.

This report is intended as a general discussion of the
problem of the use of such scenarios in system evaluation.
Specifically addressed are the issues of the definition and

characterization of the scenario problem, the proper selection
of scenarios, the evaluation of scenario probabilities, and
the use of scenarios in both system design and in choosing
among a specified set of alternative systems.

TABLE OF CONTENTS

FIGURES

ACKNOWLEDGMENT

ON USING SCENARIOS IN THE EVALUATION
OF COMPLEX ALTERNATIVES

1.0   INTRODUCTION

This report presents some ideas about scenario usage in
making complex decisions. The specific examples we have in
mind consist for the most part of evaluations of large mili-
tary systems, but the ideas are far more general than that;
we hope they apply to any complex evaluation context, military
or civilian. The key feature that makes scenarios necessary
in such situations, of course, is that the systems or what-
ever being evaluated will operate in a highly uncertain
future.

The first step in evaluating military systems (or other
social responses to anticipated future problems) is evaluation
of the need for a response. This prior assessment requires
forecasting the nature of the problem, for example, the
intentions and capabilities of potential U.S. adversaries.
If existing or projected U.S. capabilities do not meet
anticipated future needs, a social response to the resulting
deficiency may be necessary.

Given a projected deficiency, performance requirements
for systems or procedures for remedying it lead to alterna-
tive conceptual designs. If the analysis justifies a system
procurement, the steps leading to such a procurement are
initiated. These steps will usually lead to several alterna-
tive proposed systems or other solutions, each of which must
be evaluated. Procedures for doing so vary from global

1

judgments made by committees to simulation to utility-type
modeling. Whatever the procedures, the resulting evaluations
will be conditional on the nature of the problem or sequence
of events being considered. That is, they will depend on
implicit or explicit scenarios.

Our point is that every step in this process, from
initial recognition of the mismatch between problem and
capability to ultimate evaluation, not only of the proposed
solution, but of its delivered realization, must depend on
hypothetical statements about the future, that is, on sce-
narios. This paper first discusses the underlying conceptual
problem, known in decision analysis as the bushy-mess problem.
Then it examines what we mean by scenarios. Finally, it
spells out ideas about how to use scenarios both for design
and for choice among alternative designs.

2.0 THE BUSHY-MESS PROBLEM OF DECISION ANALYSIS

In principle, any decision is a straightforward matter
from a decision-analytic point of view. The action options
are defined. Contingencies not under the control of the
decision maker that might influence the merits of each
available option are specified. For each possible option
and each contingency relevant to it, a number called a
utility is found. This utility represents the attractive-
ness of the option if that contingency occurs. Sometimes it
is appropriate to recognize that subsequent actions will
occur and that their attractiveness will also be influenced
by subsequent contingencies, or uncertain events. The
result is a tree structure, typically drawn as growing from
left to right. Associated with each twig at the right-hand
edge of the tree, a utility representing the desirability of
that particular sequence of acts and events must be specified.
For each environmental contingency, a probability must be
specified; information-gathering that helps to specify such
probabilities may be included in the decision tree. Given
the tree structure, the utilities, and the probabilities, a
straightforward algorithm permits determining which of the
available options has the highest expected utility and,
therefore, should be chosen.

This solution-in-principle to decision problems too
often has a major practical flaw; it is too complicated to
carry out. The decision tree can quite easily become a
large bushy mess (to use Howard Raiffa's metaphor). While
computers can easily handle the sheer bulk of bushy messes,
the real difficulty of bushy messes is not in processing

them, but in eliciting the attendant utilities and probabilities. Such numbers are typically human judgments, and the cost and complexity of eliciting even as few as 10,000 such judgments (a small number compared to the number required by many bushy messes) is prohibitive.

Moreover, most decision makers feel that it should not take a 10,000-number input to make a choice among, say, five alternative courses of action. Indeed, decision-making can be regarded as an information-destroying process. Large numbers of bits of information (in the information-theoretic sense) enter the decision problem; very few emerge at the decision point. It must be and, indeed, is the case that many of those bits of information, many of those 10,000 or more numbers, are virtually or completely irrelevant to the decision. Why not, then, confine attention to the important small set of numbers, say 100 or less, that are really meaningful in the decision?

The difficulty, of course, is one of knowing which 100 numbers are relevant, that is, of pruning the decision tree (to borrow another Raiffa metaphor). How can one know which actions to eliminate without analysis, which environmental contingencies to ignore, which utilities and probabilities not to assess? While we know of no systematic answer to this question, which lies close to the heart of the artistic side of decision analysis, we would like to provide a few suggestions and a vocabulary for talking about one approach to the bushy-mess problem: the use of scenarios.

2.1 How Scenarios Help Reduce Bushy Messes

The essential idea of a scenario is that it should not be necessary to consider all possible environmental con-

tingencies on which the utility of an action may depend in order to assess fairly well the expected utility of that action. A subset, perhaps a small subset, of these contingencies may be enough. Suppose a full decision analysis would require analysis of a tree consisting of six layers, three successive actions each followed by its possible outcomes. Suppose, for arithmetical convenience, that at each action node there are five options and at each event node there are five alternative events, in this case, possible outcomes of the actions. The total number of twigs at the end of the tree is 5^6 = 15,625. In addition to assessing that number of utilities, it will potentially be necessary to assess the 16,275 relevant probabilities (or 13,020 if additivity constraints are taken into account). Although internal structure would surely reduce the problem to less than 28,645 independent assessments, the chances of reducing it to reasonable size seem slim. On the other hand, suppose that one considers only two outcomes at the first event node and only one at each of the other two. Even retaining five options at each action node (unlikely, for this simplifying approach), the total number of utilities to be assessed becomes 250 and of probabilities becomes 2. Moreover, it is extremely unlikely that 252 assessments will be necessary; the structure will be simple enough so that many portions of it can be eliminated by inspection without explicit utility assessment.

Obviously, one cannot be certain that the decision that seems best from inspecting such a pruned tree would also seem best if the whole tree were analyzed. The key lies in the pruning procedure. One would like the unpruned branches to represent the pruned ones in some sense so that the

5

expected utility calculated for the five actions being considered would be close to what would have been calculated by a full analysis.

2.2 Contexts for this Discussion

The preceding definition of the scenario problem is general to all decision analyses, but since our thinking has been influenced primarily by military examples, and we shall confine our attention to them. Most of these analyses have been evaluations of alternative system designs for future large military systems, none of which will be deployed for many years. Since the future circumstances bearing on their usefulness are complex and uncertain, the set of future contingencies, if not infinite, is at least unthinkably large. Detailed decision analysis would be proposterous. Yet explicit consideration of the major environmental contingencies on which system effectiveness depends is necessary for any meaningful evaluation of system designs.

The problem of complexity becomes still worse if one thinks of developing, rather than evaluating, a system design. A large military system may embody many thousands of performance parameters, each variable over a significant range. Trade-offs of one aspect of performance against another and of all against cost may enter into any single system design problem. Thus, instead of five initial action options, there may be tens of thousands interrelated by complex sets of constraints. But in system design as in the evaluation of alternative system designs, explicit consideration of future environmental contingencies is crucial.

Most of what has been said in the previous two paragraphs about military systems applies equally well to major

6

social systems.   Although our examples come from military
contexts, the arguments which they support apply equally
well to non-military problems.   Any argument for using
scenarios to reduce the complexity of a decision problem in
a military context serves as an argument to reduce such
complexity in a social context, for any decision requires
explicit consideration of alternative futures.

3.0 STAGE SETTING, ACTION, AND SCRIPT GENERATION

3.1 What is a Scenario?

For purposes of this paper, it would be desirable to precisely define the term "scenario," for there exist numerous meanings for the term, and the possibility of miscommunication is a potential problem. In Section 4.1, we provide a definition that allows us to proceed with the rather abstract discussion of the problem. However, a specific definition of the term will depend on the use to which scenarios are put, and it is just that use that we are addressing in this paper. In fact, an ambitious goal for the paper would be a set of universal characteristics that define scenarios as used in evaluation and design problems. That ambitious goal is beyond our reach at present. Instead, we shall proceed by example and try to rule out what we do not mean while at the same time attempting to clarify the term. A possible definition of the term "scenario" is the following: a hypothetical sequence of events usually set in the future. All the definitions and instances we are familiar with fit this description. But how detailed is that sequence of events?

This question is a bit like the familiar question of how much information to include in an article or report. The standard cliche is that a report should be long enough to cover the subject and short enough to be interesting. The same advice, applied to scenarios, would make them quite long, almost certainly too long for practical purposes. We have previously suggested that a scenario is equivalent to one path through a decision tree. But decision trees can vary in complexity, in the number of act and event nodes

8

included. A tree, or a scenario, comprising a single branch
of a tree, detailed enough to cover all the relevant facts
is almost certainly too detailed for easy use in evaluation.
Both increased length and a corresponding convincingness
inversely proportional to probability often make scenarios
too unwieldy (and often misleading) for practical purposes.

Moreover, there is an obvious trade-off between the
amount of detail of each scenario and the number of scenarios
that one can afford to generate and consider. One tradition,
embodied in war games, case histories, and many similar
contexts, is to make scenarios very specific and consider
very few of them, often only one. We oppose that tradition
in evaluation contexts, though it may make sense in training
contexts. We believe that more sketchily defined, far
briefer scenarios can often capture all that is crucial to
making discriminations among the alternatives of a decision.
For example, in evaluating the relative value of having the
technical effectiveness associated with particular aircraft
weapons systems, the factor of overriding importance may be
the degree of involvement of Soviet technology, which can be
characterized by only a few details. A scenario for evaluating
such systems can thus be sketchily defined, and scenarios
that are sketchily defined can be more easily generated, and
more of them can be used, if we can assume, of course, that
necessary probability and utility judgments can be made with
respect to them. We shall characterize potential scenarios
as having specified levels on some subset of an infinite set
of dimensions (or variables). The specified levels of these
dimensions would be the scenario details. For example, in
the evaluation of the Single Channel Ground and Air Radio
System (SINCGARS)[1] for the U.S. Army, three scenario dimensions

_____

[1] James O. Chinnis et al., <u>Single Channel Ground and Airborne
Radio System (SINCGARS) Evaluation Model</u>, Technical Report
DT/TR 75-2 (McLean, Va.: Decisions and Designs, Inc. 1975).

9

were varied. These were: the existence or lack of a definable frontal edge to the battle area (FEBA versus non-FEBA), type of terrain (mountains, forests, open country, or urban areas), and radio platform (man, ground vehicle, or aircraft). Descriptions of the levels of these dimensions were provided. From the levels of these three dimensions, twenty-four (2x4x3) scenarios were created for use in the evaluation. A scenario consisted of three details, that is, the specifications of the levels of each of these dimensions. Other details were implicit in each of these scenarios. Examples are the fact that the radio was being used by U.S. Army personnel in combat operations in the specified setting, and it was assumed that weather was not a factor in radio performance.

This SINCGARS example illustrates a fact about scenarios. No matter how detailed the description provided by a scenario, other details can be added. Often the sketchy set of scenario details such as that of the SINCGARS example is embedded in a larger set of other scenario details that are provided for clarity and verisimilitude and that aid in the relevant probability and utility assessment tasks. A potential scenario can, therefore, be thought of as consisting of two parts. One part is the set of details that specify the levels of variable scenario dimensions that change from scenario to scenario. These details are associated with the major utility differences among alternatives to be evaluated in the scenarios. A second part of a scenario consists of details added for clarity and verisimilitude to facilitate probability or utility assessments. These details may be the same for all scenarios, and if so, the analysis is anchored to a subspace of all potential scenarios. Alternatively, these details added for clarity may vary across scenarios. An example would be the following: an important

detail in the assessment of the utility of the technical
capability of alternative naval aircraft weapons systems
involves the area of the world in which the systems will be
deployed. The value of that capability can be different in
operations in the Middle East as opposed to operations in
the South Atlantic. The types of threats are different for
these areas as are the potential values of the outcomes of
the operations. These area descriptions may suffice for
capturing the differences in the value of the technical
capability of the air systems, but it may be the case that
the operational expert assessing utilities requires a more
specific setting, say, Lebanon as opposed to Venezuela.
These area details can be added to the scenarios for clarity.
In addition, certain details such as the size of the fleet
in terms of number of aircraft carriers and associated
numbers of squadrons of aircraft can be fixed across scenarios.
Note also that the explicit set of scenario details also
implies an implicit set of details to which the expert may
attend while making judgments.

In this paper, we shall emphasize the set of details
that change from scenario to scenario. The criteria for
effective scenarios will determine the size of that set.
However, we cannot and shall not ignore either the constant
or the implicit details of scenarios, for the choice of
specific sets of details for emphasis in evaluation is
necessarily based on a set of important assumptions, the
nature of which must be clarified. The nature of these
assumptions as well as criteria for choice of scenario
details will be discussed later in appropriate sections of
the paper.

3.2 The Parts of a Scenario

In contexts like that of the SINCGARS evaluation, we find it helpful to think of the path through the decision tree that we are exploring as divided into three parts: stage setting, action, and script generation. The stage setting corresponds to what might be called scenario generation in, for instance, a war game. It specifies the operational setting and all relevant actions and events prior to the deployment of the system (or whatever we are evaluating). Note that the stage setting will typically include "acts" defined in the traditional design-theoretical sense as occurrences that the decision maker controls, as well as events not under his control. Since a single stage setting does not consider alternatives to the acts specified in it, both such antecedent acts and the events are alike considered as events.

Embedded in the stage setting will typically be particular values of the variables that we have chosen to define a scenario. These will be called conditioning variables since we expect the utility of whatever is being evaluated to be conditional on them. Thus, the stage setting consists of a scenario, that is, a particular set of values of the conditioning variables, plus whatever additional detail need be added for clarity and verisimilitude.

Action simply means occurrence of whatever we are evaluating. In the context of evaluating a military system, for example, it would mean deployment or use of that system.

Script generation means specification of a set of guesses about the consequences of the action, for a given stage setting. A script in effect plays out the acts and

events subsequent to the act actually being evaluated to the point at which a utility can be assigned to the sequence of stage setting, action, and script generation. While it would obviously be possible and may often be appropriate to consider alternative sequences of actions and events subsequent to the evaluation of the act, we have most commonly encountered situations in which one set of events subsequent to the act was enough; the real uncertainties occurred in stage setting. If a single script is used to evaluate an act for a given stage setting, the conditional utility of the act is simply that assigned to that script. If more than one script is used, probabilities must be assigned to each and the conditional utility is calculated by using the assigned probabilities. Assessment of such probabilities is not different from assessing the probabilities involved in stage setting. Both topics are examined later in this paper.

Figure 3-1 illustrates the three parts of a scenario in the SINCGARS example.

3.3 The Analogy to Stratified Sampling

Obviously the stage setting and script generation processes represent enormous reductions in the size of the decision tree and, thus, enormous simplifications of an evaluation. Whether or not they are exhaustive with respect to the conditioning variables that have been chosen for particular attention, they held all other variables constant. This approach has a partial, though by no means complete, similarity to stratified sampling. A stratified sampling process exhaustively partitions the universe to be sampled and uses auxiliary information to specify the relative incidence of each stratum in that universe. Such a process

STAGE SETTING: *Prior Acts and Events*

ACTION: *Introduce System*

SCRIPT GENERATION: *Subsequent Acts and Events*

EVENT 1:
BATTLE
AREA EDGE

EVENT 2:
OPERATING
ENVIRONMENT

EVENT 3:
PLATFORM

FEBA

NON-FEBA

MOUNTAINS

FORESTS

OPEN
COUNTRY

URBAN
AREA

MANPACK

VEHICLE

AIR
SYSTEM

DEPLOY SYSTEM A

DEPLOY SYSTEM B

DEPLOY SYSTEM C

DEPLOY SYSTEM D

Figure 3-1

HYPOTHETICAL DECISION TREE FOR THE SINGGARS EVALUATION

corresponds to the selection of conditioning variables and the assessment of their probabilities. A stratified sampling process then samples randomly within each stratum by using stratum relative frequencies to extrapolate to the population. Nothing analogous to random sampling occurs within scenario usage; it operates as though each stratum in a stratified sample were represented by a single individual.

The key point here is that, in selecting conditioning variables, as in selecting strata for a stratified sample, one attempts to represent exhaustively the universe being examined. The sense can be carefully defined. In stratified sampling, one tries to pick strata that will be different with respect to the variable(s) being measured, preferably widely different. And one defines the strata so that collectively they exhaust the universe of interest. Similarly, in scenario selection one tries to choose conditioning variables that will affect as heavily as possible the utility of the system, action, or whatever being evaluated. Unless one anticipates that at least some of the utilities associated with one or more of the acts being evaluated will differ markedly from one state of the conditioning variable to another, the choice of conditioning variable is a poor one.

If insight into the problem permits, one can be even more specific. The choice of conditioning variables should help to differentiate among the acts being evaluated. A conditioning variable that affects all acts being considered in the same way is not useful for evaluation even if the size of the effect is large. On the other hand, small effects are clearly not of interest either. So the ideal conditioning variables will affect utility both strongly and markedly, that is, in a way that differentiates among acts

being evaluated. (Again, as indicated, conditioning variables that affect all acts under consideration in the same manner may be included in scenario descriptions for clarity and verisimilitude and as judgmental anchors.)

4.0 PROCEDURES FOR SCENARIO USE IN EVALUATION
FOR DESIGN AND CHOICE

4.1 Terminology

We start by defining language and some notation.

4.1.1 Scenarios - Scenario space, S, the set of potentially relevant scenarios, is composed of individual scenarios, s_i; each is multidimensional, that is: $S = \bigcup_i s_i$, where $S = (s_{i_1}, s_{i_2} \ldots s_{i_m})$. Some subset of S will actually be used in evaluation. Establishing the dimensions of S to be used for evaluation is a major part of the design or choice problem. For example, in determining the utility to be attached to the technical capabilities of different air systems against different threats, how should scenarios be defined? What are the relevant dimensions of scenario space? The reader may think of specification of a particular scenario detail as earlier discussed as the setting of some level of a dimension of scenario space. For example the specification of an area of the world, say the Middle East, would be the specification of the scenario dimension corresponding to world area. As another example, a possible partition of the world could use the following dimensions:

o   Areas of the world,

o   Type of U.S. military action, and

o   Level of Soviet technology faced.

These dimensions can be characterized by the following levels:

17

o   Areas of the world - North Atlantic,
        South Atlantic, West Pacific, East Pacific,
        Indian Ocean, and Mediterranean Sea. (These
        six areas are so defined as to partition the
        entire world.);

o   Type of U.S. military action (e.g., general
        deterrence, crisis, unilateral military
        action, NATO War, and other); and

o   Level of Soviet technology faced (e.g.,
        Non-Soviet weapons, unsophisticated Soviet
        weapons manned by non-Soviet personnel,
        sophisticated Soviet weapons manned by
        non-Soviet personnel, and sophisticated
        Soviet weapons manned by Soviet personnel).

A possible scenario could then be a unilateral
military action by the U.S. in the Mediterranean area in
which Soviet weapons manned by non-Soviet personnel. This
scenario is fairly specific but can be made even more specific
by specifying that the action occurs in Lebanon against
specific numbers and kinds of Soviet weapons. This would
involve redefinition of scenario dimensions or further
refinement of dimension levels or both.

The specification of the dimensionality of a
scenario is, in essence, the specification of the minimum
amount of detail with which a scenario can be described to
allow valid judgments of the type necessary for the evalua-
tion. In essence, as earlier discussed, S can be thought of
as being of infinite dimensionality. However, many of the
potential details will be ignored. The actual dimensions
chosen to characterize scenarios will consist of those that

vary from scenario to scenario, and those that are constant--that is, the details are the same--from scenario to scenario. The number of such dimensions is the dimensionality of S.

4.1.2 <u>Options</u> - The set of potential system designs will be denoted as O, the option space, a potentially infinite set of arbitrary dimensionality. Each option o_j will be a multi-attributed alternative with values on each of the n attributes that constitute the dimensions of O. That is, $O = \bigcup_i o_j$ where $o_j = (o_{j_1}, o_{j_2} \ldots o_{j_n})$.

Attributes are system characteristics under the control of the system designer, that determine the performance of the system in a particular scenario, and can be described independently of any scenario. Consider, for example an electronic warfare (EW) suite aboard a particular naval vessel faced by an incoming threat (missile or air system) with an active seeker looking for the vessel. Of interest would be the ability of the EW system to handle the seeker whether by decoying it, deceiving it, or jamming it. These abilities are functions of several system characteristics. For example, the ability of the EW system to jam the seeker is a function of such system characteristics as output power, antenna directivity, modulation capability, frequency coverage, and the like. As such, these characteristics are more specific than jammer ability, and they are also functions of other, more highly specific design characteristics. For example, the output power is a function of the size and number of transmitter tubes. The question then becomes one of defining exactly what an attribute is. An attribute is a system characteristic that can be manipulated by design modifications. As such, attributes can be as general as ability to decoy, ability to jam, and the like, or as specific as the resistance of a particular resistor in a power supply.

It is helpful to view the designer as having a
large number of dials under his control, one for each of the
attributes of the system. A particular system design corre-
sponds to a set of specific dial settings. We will call
moving a dial to alter a design "tweaking."[1] There are two
important points to be made about tweaking. First, tweaking
with respect to a particular attribute potentially corresponds
to the manipulation of many different system character-
istics, and the relationships of actual design combinations
to any particular attribute dial setting are a many-to-one
relationships. That is, depending on attribute definition,
many different design combinations can produce the same
level of an attribute. The more general the attribute, the
more likely this is to be true. "Tweaking" describes the
altering of a detailed system design at whatever level one
chooses to discuss, and tweaking with respect to a particular
attribute will be discussed as the manipulation of a single
dial.

A more important point about tweaking is that
it is intolerably multi-dimensional. That is, it is often
impossible to tweak with respect to a single attribute.
Design constraints are such that certain dials must be
simultaneously turned, and the settings of some dials can
affect the settings of others. This corresponds to the
"environmental correlation" problem addressed in discussions
of evaluation of multi-attributed alternatives.[2] Put very

---

[1] We borrow this term from electronics slang. Electronics
technicians often speak of fine tuning or tweaking a system
for best performance.

[2] See Ward Edwards, How to Use Multi-Attribute Utility Measurement
for Social Decision-Making, Technical Report 001597-1-T (McLean,
Va.: Decisions and Designs, Inc., August, 1976).

simply, environmental constraints limit the number and
characteristics of available designs for evaluation. That
is, the process of system design which will be abstractly
discussed is very complex, and the location of designs
requires the simultaneous consideration of many non-inde-
pendent system attributes. The actual process of tweaking
can, therefore, be greatly aided by computerized procedures
that take into account the constraints on tweaking and can
automatically modify design characteristics in response to
these constraints. The reader familiar with the techniques
and problem areas of the general field of operations research
will note that optimal system design can be viewed as a
constrained optimization problem. As such, the problem has
received in-depth attention. For further discussion of this
aspect of the problem, we therefore direct the reader to
that literature,[3] and having acknowledged this complication
with tweaking, we will hereafter ignore it.

4.1.3 Utility - The expected utility of a particular
option is a function of the performance of the option in the
scenarios chosen for evaluation of options. Performance is
a very general term. System evaluation involves many different
factors such as performance in defeating threats, operational
acceptability, adaptability to interface with future systems,
life-cycle cost, and the like. Depending on the choice of
scenarios, some of these aspects will be scenario independent,
others not. For example, certain aspects of operational
acceptability can be viewed as being independent of combat
scenarios. Examples for a general weapon system might be
personnel manning requirements, space and volume utilization,

_____

[3] See, for example, Harvey M. Wagner, Principles of Operations
Research: With Applications to Managerial Decisions
(Englewood Cliffs: Prentice Hall, 1969).

safety, skill level required to operate, training needs, and functional modularity. However, other aspects of operational acceptability are clearly not scenario independent.

Performance will be denoted by an r-dimensional space P. A particular option o_j in a particular scenario s will achieve a performance denoted by the r dimensional vector $p_{ij} = (p_{ij_1}, p_{ij_2} \cdots p_{ij_r})$. The choice of performance dimensions depends on the system and context, but obviously the evaluation must have some basis (perhaps subjective) for assessing each performance dimension for each option, o_j, though not necessarily for each scenario, since some performance dimensions may be irrelevant to some scenarios.

The dimensions of P could correspond to the attributes of O. For example, an EW system could be described in a rather general sense by such system characteristics as ability to jam an active seeker, ability to decoy an active seeker, ability to deceive an active seeker, and the like. The dimensionality of P would then be the same as that of O. On the other hand, EW system performance could be described in terms of such characteristics as the number of incoming missiles defeated, ability of the system to defeat a specified combination of threats, or some other related measure. More often we prefer the latter approach to the former.

As indicated, the ultimate goal of evaluation for design or choice is the determination of expected system utility. Thus the final space to be discussed is U, the utility space. U will be defined on the vector product of the O, S, and P spaces. The utility that is assigned to an option is a function of the performance of that option in a particular scenario. The utility of the performance p_{ij} of o_j in s_i will be denoted as $u(p_{ij})$. The utility space U can

also be analyzed into its appropriate dimensions by considering the respective dimensional utilities of the performance, that is $u_k(p_{ij_k})$. Then the overall utility $U(p_{ij})$ is some function F of these dimensional utilities.

$$U(p_{ij}) = F[u_1(p_{ij_1}), u_2(p_{ij_2})...u_k(p_{ij_k})...u_r(p_{ij_r})]$$

At this point we need to consider the scales of measurement used for utility. Since utilities are usually considered to be measured up to a linear transformation, both origin and unit of measurement need to be specified. Our ideas about these matters are fairly conventional, and have been treated (from somewhat different points of view) by Edwards (1976) and by Keeney and Raiffa (1976). The essential point is that the units of measurement of the utilities of different dimensions of performance either must trade off directly against one another or else must be weighted appropriately to make such trade-offs meaningful. The references cited above, both oriented toward weighting, proposed methods for ensuring that this will occur. It is often convenient to treat really good performance on a dimension of performance as having a utility of, say, 100, and really bad performance as having a utility of 0. We shall adopt this convention.

An important utility is that of the performance of the current system. The current system will be denoted as o_c. Thus,

$$o_c = (o_{c_1}, o_{c_2}...o_{c_n})$$

The performance of the current system in scenario s_i will be called p_{ic}, where $p_{ic} = (p_{ic_1}, p_{ic_2}...p_{ic_r})$. The utility attached to the performance of the current system in scenario s_i could be written as

23

$$u_{ic} = u(p_{ic}) = F[u_1(p_{ic_1}),\ u_2(p_{ic_2})\ldots u_k(p_{ic_k})\ldots u_r(p_{ic_r})],$$

but for notational simplicity we shall instead write

$$u_{ic} = F(u_{ic_1},\ u_{ic_2}\ldots u_{ic_r}).$$

The utility of the current system can be used as a sort of baseline against which other systems can be compared within scenarios. Indeed, scenarios can be evaluated in this way. A high-probability scenario in which the current system performs miserably is obviously important, since it defines a problem in need of solution. (A high-probability scenario in which the current system performs excellently but a proposed future system would perform miserably would also be important, but we shall make the obviously optimistic assumption that such cases will not arise and so shall ignore them.)

Slightly formalizing the preceding paragraph, we shall define I_i, the importance of the ith scenario, as follows:

$$I_i = k\ p_i/u_{ic},$$

where k is a scaling constant.

For a rough first approximation, I_i can be used as an index of whether or not a new system is needed. If all values of I are relatively low, either because the probability of scenarios in which the current system performs miserably is low or because the quality of performance of the current system in all scenarios is relatively good, then a new system need not be considered. Only if relatively high values of I_i can be found is it worth while to consider replacing the existing system.

24

4.2 Utility for Options as a Function of Scenarios: Kansas vs. Idaho

4.2.1 <u>Multiple scenarios: one system</u> - Consider first an analysis to determine whether or not a new or modified system design is required. Assume that a representative set of six scenarios has been created. o_c is projected into these scenarios, and u_{ic} is assessed for each scenario s_i. Examples of possible results appear in Figure 4-1a through 4-1c. In Figure 4-1a, o_c has acceptable utility in all six scenarios, and utility is fairly invariant as a function of scenarios. Since the performance of o_c is acceptable in Figure 4-1a, no new system is necessary. In case 4-1b, o_c performs miserably. The possibility that a new system might do better certainly should be explored. Of course, nothing in this analysis guarantees that a better system can be found. However, if we assume that the six scenarios in case 4-1b are different from those in case 4-1a, we should examine S for purposes of ascertaining the source of the relatively constant difference in performance. Is there a scenario detail or group of details that is associated with this difference, and if so, what are the design implications? In case 4-1c, the answer is uncertain. Perhaps the variation from scenario to scenario simply represents the attractiveness of the scenarios themselves, independently of what the system can accomplish within them. Or perhaps a new system could do better for those scenarios within which o_c does badly.[4] Obviously, the potential of a new system should be explored.

---

[4] It would be possible to eliminate from evaluative consideration the attractiveness or unattractiveness of the situations represented by the scenarios themselves, which is after all irrelevant to the evaluative process, by using some version
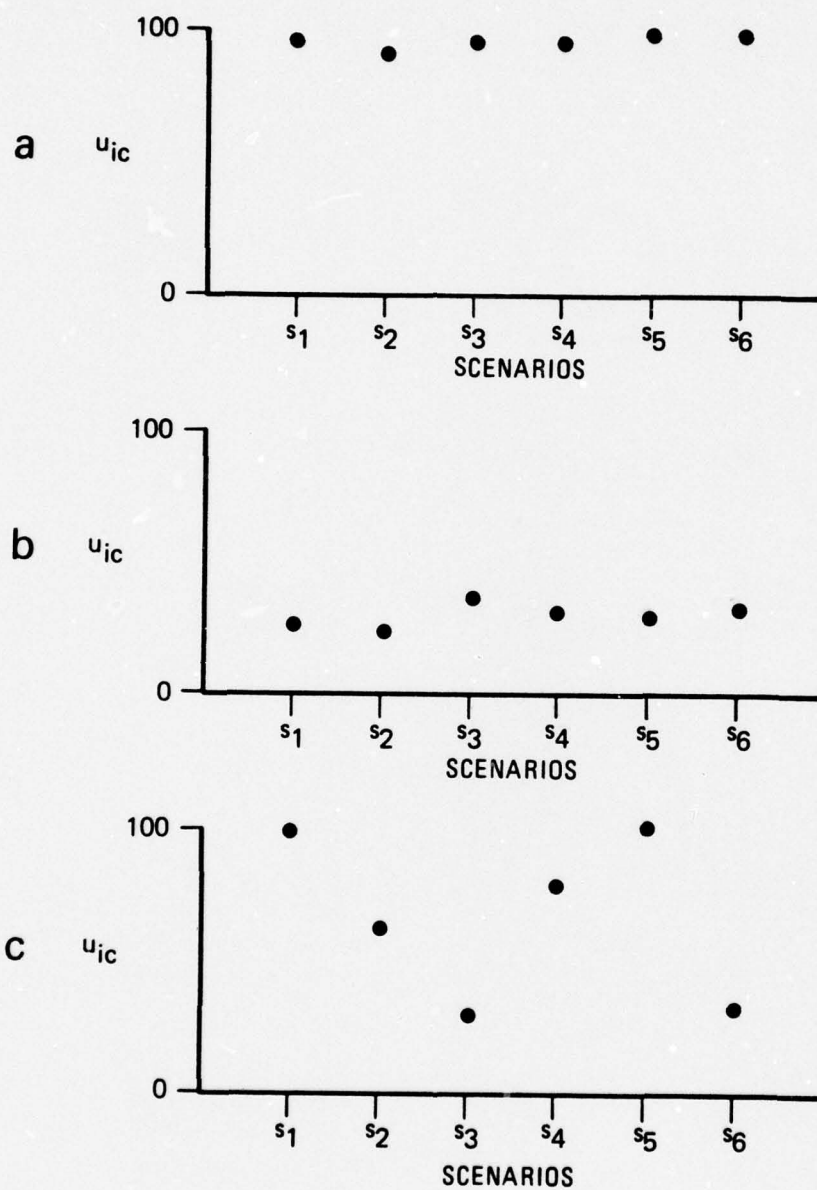
Figure 4-1

POSSIBLE RESULTS OF EVALUATING THE
UTILITY OF OPTIONS o_c IN s_1 -s_6

The situation in which the utility for an option is fairly invariant as a function of scenarios, as in Figures 4-1a and 4-1b, will be referred to as being in the "plains of Kansas" with respect to that option. Obviously, if it were known a priori that the utility assessment situation was of a Kansas type, any of the six scenarios could be used for assessing the expected utility of o_c. The situation in which utility is highly variable as a function of scenarios will be described as being in the "mountains of Idaho." In the Idaho case exemplified by Figure 4-1c, the determination of the expected utility of o_c is a more complicated problem necessitating investigating o_c in more scenarios. When the situation is of an Idaho type, the examination of scenario space must involve greater precision in order to ascertain expected system utility. Note that precision can be obtained by increasing the number of scenarios or increasing the specificity of the scenarios.

Suppose, for example, that the only difference in s_1 through s_6 in Figure 4-1 is the area of the world as

---

[4](cont.) of the concept that L.J. Savage invented and called loss, but that is for the most part known in decision analysis as regret. Essentially, this involves defining the best possible performance of any system within a scenario as zero for that scenario, and measuring within-scenario differences between other performances and that one. But the technical difficulties are formidable. Presumably it would be necessary to define the zero point separately for each dimension of performance. In situations in which the action space is ill-defined or incompletely specified, such as system design contexts, it might be very difficult to know the best possible performance until after the process was complete and the knowledge no longer relevant. Much the same comments apply to defining the worst possible performance within each scenario as zero for that scenario.

described in the example in Section 4.1.1 on naval aviation.
Suppose those areas for the six scenarios are as follows:

s_1- South Atlantic
s_2 - North Atlantic
s_3 - Mediterranean
s_4 - Indian Ocean
s_5 - Eastern Pacific
s_6 - Western Pacific

In Figures 4-1a and 4-1b, the o_c has about the same utility in any area, and the expected utility could be estimated using any of the scenarios. As indicated, such is not the case in Figure 4-1c. Suppose that the six scenarios are reduced to two. Scenario A will involve the South Atlantic, North Atlantic, and Mediterranean, that is, s_1, s_2, and s_3. Scenario B will involve the Eastern Pacific, Western Pacific, and the Indian Ocean, that is s_4, s_5, and s_6. This reduction in specificity of scenarios would have little effect in Figures 4-1a and 4-1b, but could have quite an effect in Figure 4-1c, depending, of course, on the respective probabilities of s_1-s_6.

4.2.2 Multiple systems: one scenario - Assume that the problem of scenario identification has been solved and s_1 has been identified as a very important scenario to consider. The utility of several other options can be examined in s_1 as illustrated in Figure 4-2.

Because s_1 is an important scenario, o_c has low utility in all three cases. In case 4-2a, none of the five designs improve much on o_c. In case 4-2b, all five designs
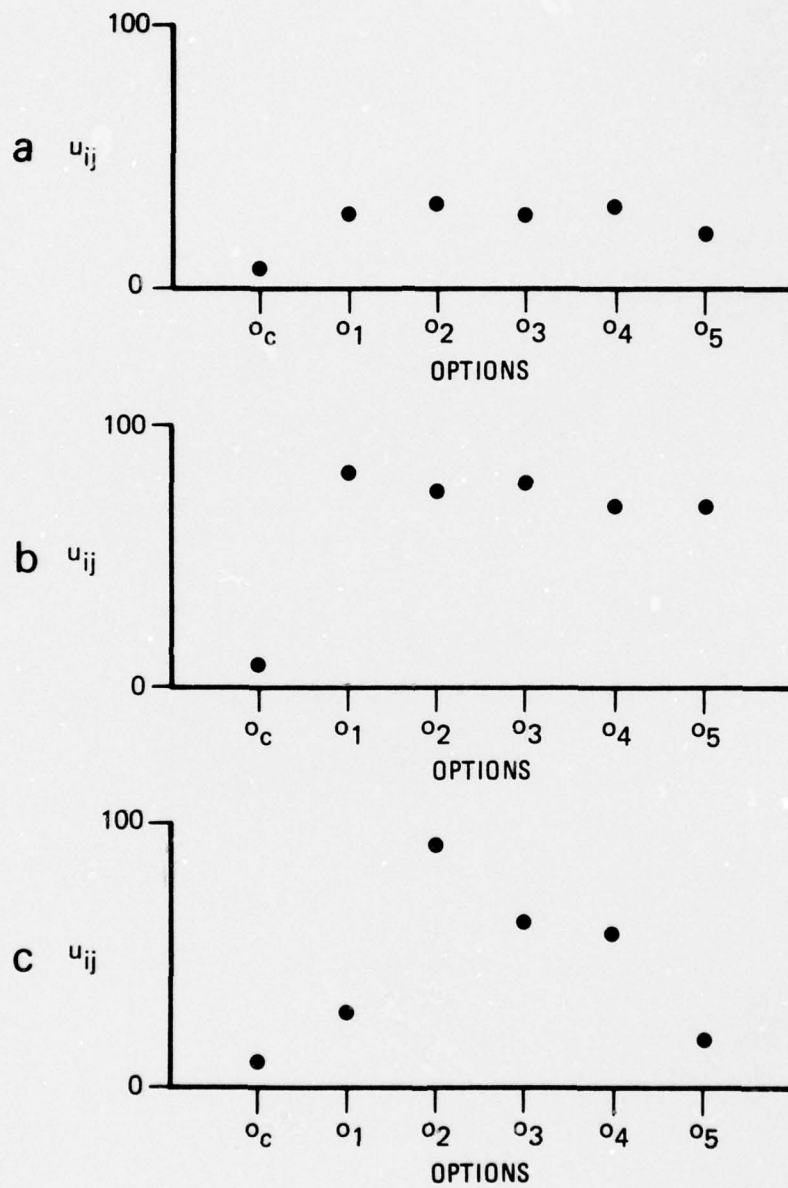
Figure 4-2

POSSIBLE RESULTS OF EVALUATING THE
UTILITY OF OPTIONS o_c AND o_1 -o_5 IN s_i

are far superior to o_c in utility. In case 4-2c, o_2 has highest utility, o_3 and o_4 are acceptable, and o_1 and o_5 offer only marginal improvement over o_c.

Why would a case like 4-2a occur? One possibility is that the low utility u_{ic} is not a function of system performance, but rather of other factors that are relevant to utility. For example, it may be the case that in a particular scenario, no matter how well a system performs, the consequences of system deployment will be unattractive. This simply means that system design cannot alter the distressing situation.

Similarly in case 4-2b, all designs are of equal utility. This result is quite conceivable, for most systems are designed to handle likely major threats, and although different technical procedures for threat neutralization may be adopted, the outcome is the same. The reason is that o_c is simply outmoded with respect to the particular threat in s_i and that all the modified designs being considered will handle that threat. Accordingly, in this particular Kansas-type problem, design differences have a minimal effect.

4.2.3 <u>Multiple systems: multiple scenarios</u> - If we examine one system in multiple scenarios, we can observe the variation in performance as a function of scenarios, but we do not know how other systems would perform in the same scenario. Similarly, if we observe multiple systems in one scenario, we can ascertain differences in the performances of the different systems, but we cannot be sure that this difference persists across scenarios. Note the similarity to the observation of main effects in an analysis-of-variance-type paradigm. If we observe multiple systems in multiple scenarios, we can simultaneously observe both aspects,

system effects and scenario effects, and we can also observe the iterations between systems and scenarios. These interactions are our major concern in system design. In Figure 4-3, o_c and two other options, o_1 and o_2, are evaluated for each of six scenarios. Case 4-3a is a Kansas-type situation for all three systems. It introduces, for the first time in this paper, the decision-theoretic concept of dominance. The strategy, or design, o_2 is said to dominate strategy or design o_1 if, for all scenarios (states of the world) being considered, o_2 is at least as good as o_1 and, for at least one scenario, o_2 is definitely better.

The kind of dominance illustrated by Figure 4-3a depends on the aggregation process by which the utilities of the various dimensions of performance are combined to obtain u_{ij}. Perhaps o_2 is very much better than o_1 on performance dimension k and o_1 is a bit better than o_2 on performance dimension l. If dimension k is considered much more important than dimension l, then the situation illustrated by 4-3a can result. Obviously, reversal of these judgments of importance would reverse the relative attractivenesses of the two options.

A stricter kind of dominance can be defined by requiring o_2 to dominate o_1 on each dimension of performance for each scenario. In cases in which this very strict kind of dominance exists, o_1 can and should be discarded from consideration without further ado. Even if the looser form of dominance exists, it is appropriate to discard o_1 from consideration without further analysis if the aggregation rule for combining the utilities of dimensions of performance is considered appropriate and not subject to reconsideration.
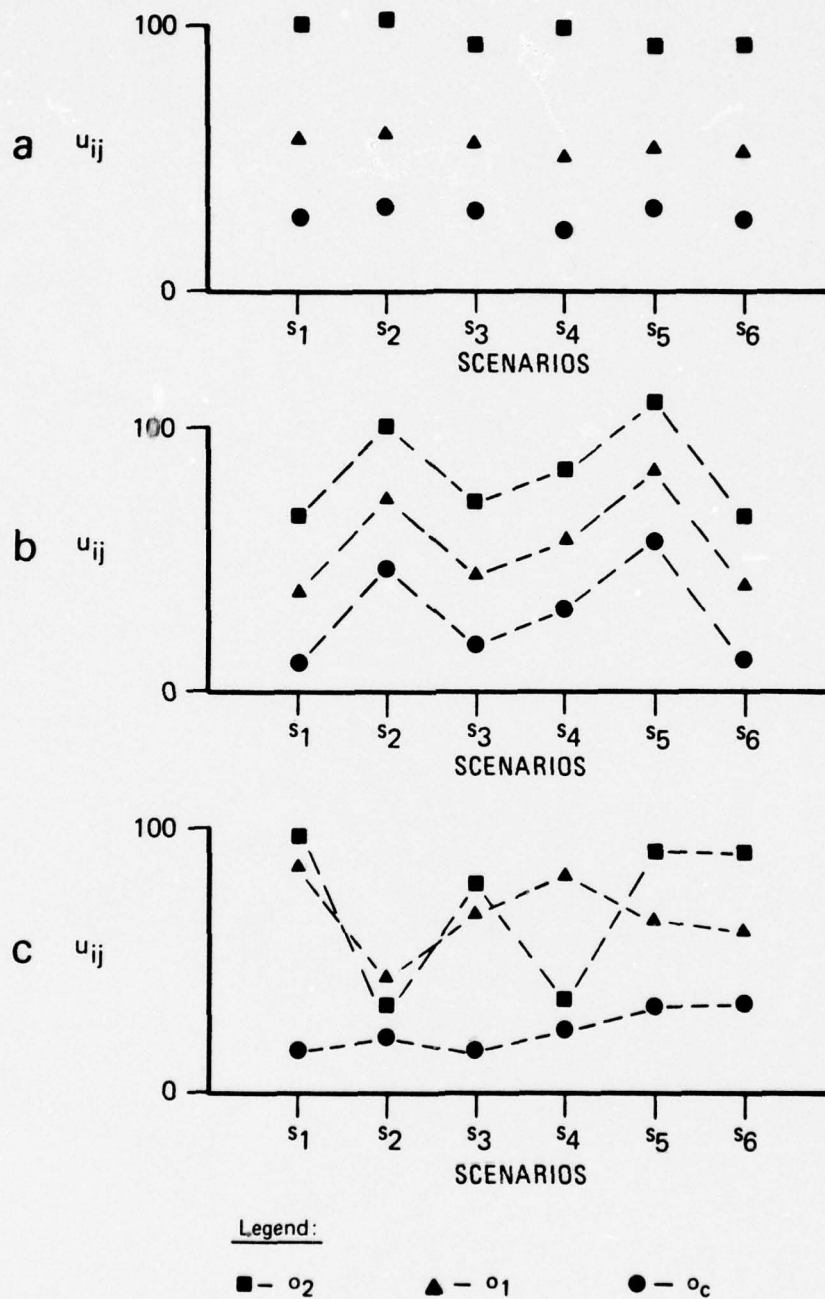
Figure 4-3

POSSIBLE RESULTS OF EVALUATING THREE
SYSTEMS o_c, o_1, AND o_2 IN s_1 -s_6

Case 4-3b is not a Kansas-type situation for any system. That is, scenarios make a difference, that difference is much the same for all systems. Presumably whatever causes good or poor performance in the scenarios is common to all systems in this case. Note that the parallel utility contours indicate that this situation which is closer to an Idaho-type than a Kansas-type situation still exhibits the kind of dominance discussed for case 4-3a.

Case 4-3c is an Idaho-type situation for o_1 and o_2 (and possibly o_c). In this situation, dominance no longer exists: there is an interaction between designs and scenarios. Option o_1 appears to be a less variable, lower risk option than is o_2 (if risk is defined as variance in utility). In design, case 4-3c is the important one. It is desirable to locate differences in options that lead to the highest utilities in scenarios. However, trade-offs between alternative designs will not occur if utility differences remain invariant across scenarios as in cases 4-3a and 4-3b. In choice, case 4-3c is the most difficult of the three. Knowing about the parallel contours in cases 4-3a and 4-3b allows heavy emphasis on one particular scenario to establish option utilities and less exhaustive treatment of other scenarios.

4.2.4 <u>Multiple dimensions of scenarios: one system</u> - As discussed, the attributes (or dimensions) of S can be viewed as scenario details. More detailed scenarios are those with information specified for a large number of attributes. The initial step in evaluating the utility of o_c in scenarios to determine the need for a new system involves examination of the attributes of S to ascertain the amount of detail necessary for establishing option utilities. Figures 4-4 and 4-5 illustrate possible results of examining two attributes of S. In Figure 4-4, scenario attribute 1

33
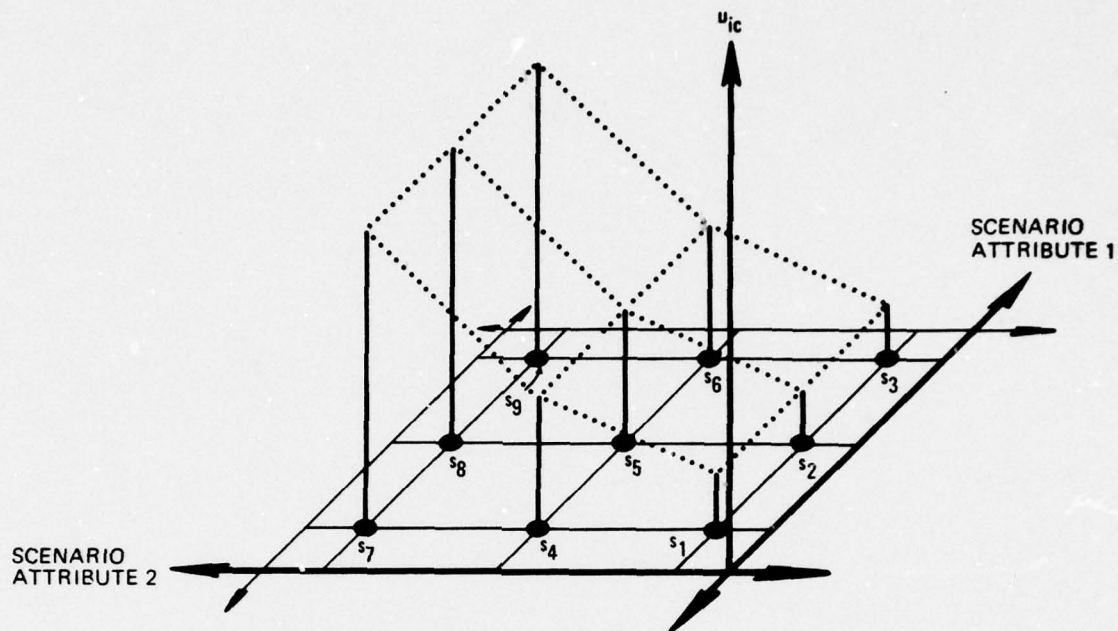
Figure 4-4

**EXAMINING THE UTILITY OF OPTION o_c AS A
FUNCTION OF TWO ATTRIBUTES OF SCENARIO SPACE**

has no relation to the utility change for o_c. For any
particular level of scenario attribute 2, utility of o_c is
essentially constant across different levels of attribute 1.
The utility of o_c is, therefore, independent of the level of
attribute 1 chosen for analysis. As will be seen later,
however, this independence does not mean that the attribute
would not be used in actual analysis requiring utility
assessments. Scenario attribute 2, on the other hand, has a
strong relation to the utility of o_c. As scenario numbers
increase along attribute 2, u_{ic} increases, being lowest in
s_1. For purposes of design, therefore, attribute 2 would be
important, and a system that could improve on the performance
of o_c in s_1 through s_3 and possibly on s_4 through s_6 would
be sought.

34

In Figure 4-5, both attributes of S are related to changes in the utility of o_c, although the relation is much stronger for attribute 2. Note that the utility of o_c is low for s_2 and s_3 as well as s_1. However, that effect is because of attribute 2, as can be ascertained by examining scenarios s_4 through s_6. In design, therefore, a system would be sought that would improve on the performance of o_c in s_1, with attention paid to both attribute 1 and attribute 2 but with greater emphasis on the latter.
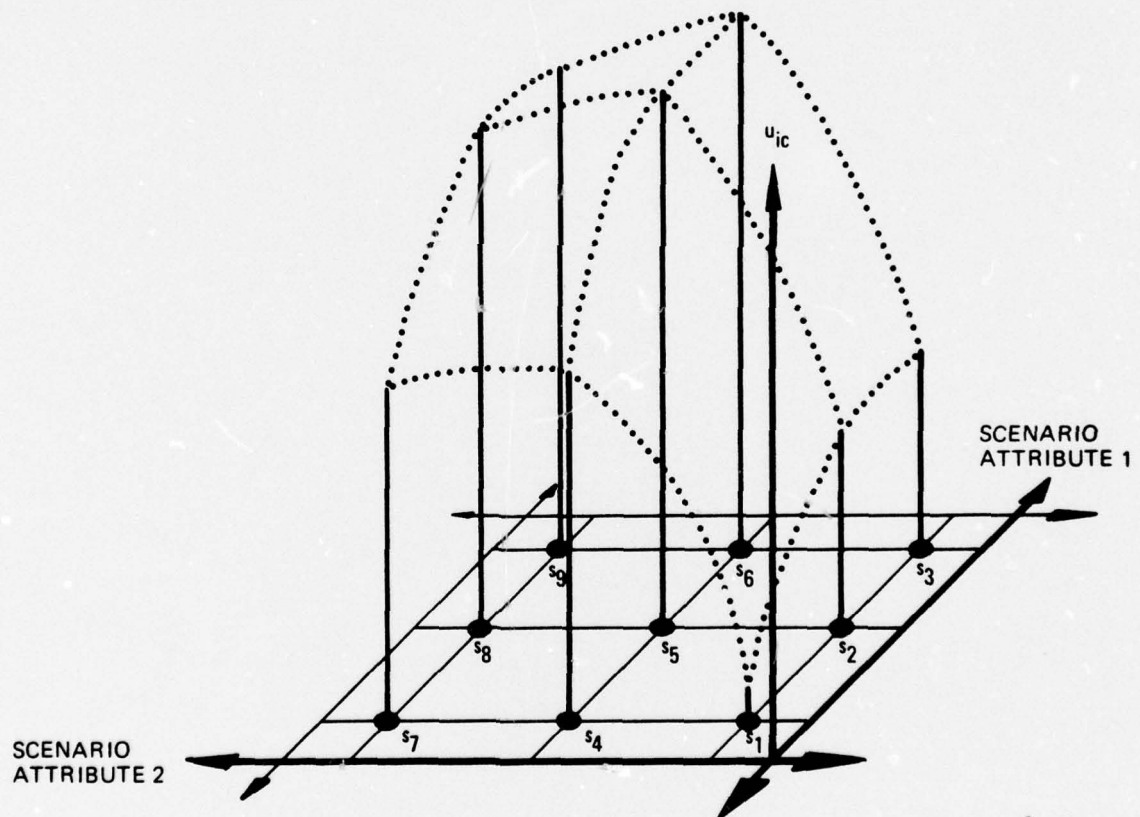
Figure 4-5

EXAMINING THE UTILITY OF OPTION o_c AS A FUNCTION OF TWO ATTRIBUTES OF SCENARIO SPACE

Continuing with this process of examining S by holding o_c constant and varying s_i, we identify the attributes (dimensions) of S that are important to determination of $u(o_c)$. Figures 4-1 through 4-5 hint at how this procedure is accomplished, but a more thorough discussion of selecting scenarios is contained in Section 4.3.

Note that in Kansas, simply because of the relative invariance of option utilities as a function of scenario, relatively few attributes will bear a strong relation to u_{ic}, whereas the opposite is true of Idaho. This fact implies that in Kansas, the choice of the s_i can proceed by using broader slices of S, that is, by using less detail than in Idaho. Because of the high variability of u_{ic} as a function of s_i in Idaho, the choice of the p_i must proceed by using highly specific scenario attributes.

4.3 How to Select Scenarios

Scenarios should ideally have three properties. First, they should be plausible or, in the decision-analytic translation of that concept, they should be relatively probable. Secondly, they should be representative, in the sense that they vary the situational dimensions that seem most important and cover appropriate ranges of each dimension varied. Third, they should be relevant to the decision problem, in the sense that they should discriminate as much as possible among the utilities of the options being considered. (Much the same principles apply to the selection of strata in stratified sampling.)

One standard approach to meeting these three kinds of needs is to make direct judgments about them. The naval air system earlier discussed is an example. The attributes

36

chosen for variation were areas of the world, types of military action, and level of Soviet technology involved. The Cartesian product of the levels of variation on these dimensions (6x5x4) would have produced 120 scenarios. Not all were used, however, since some combinations were judged preposterously unlikely. The first two of the three properties were used directly while the third (affecting utilities) was not explicitly used but entered implicitly into the selection of the conditioning variables. Since judgments of plausibility and of representativeness are relatively easy to make, they are most often used in evaluative contexts. But the real topic of interest in such contexts is the third criterion, relevance, or the ability to discriminate among utilities of the options. A plausible, representative set of scenarios that did not help to choose among options would be of little value.

One approach to maximizing utility differences is called single-threading. The idea is simply that an expert on the topic at hand can often identify very detailed circumstances in which a system, in particular o_c, will function badly and other detailed circumstances in which it will function well. Without regard for plausibility or representativeness, he can try to specify these circumstances. Then, given a set of such specifications, he can partition S in such a way as to lead to scenarios having each of such properties.

Ideally, the two approaches can be combined. Single-threading can be used to develop intuition about what variables are most relevant for inclusion in the partitioning process that defines a set of scenarios. Then direct judgments of plausibility and representativeness can enter into the detailed choice of partitions. The approach in such a

37

case involves having the judge think of peaks and valleys in Idaho (if the problem is one of an Idaho type) and relate these peaks and valleys back to dimensions of S associated with them. The nature of U as a function of S is obviously necessarily examined in doing so. Once the details of S relevant to utility discrimination have been isolated, the judge then considers the additional dimensions of S necessary for creating a representative set of scenarios. By varying details along the two sets of dimensions a representative and discriminative set of scenarios should result. But it may not be one that allows probability or utility estimation. Addition detail may be necessary for clarity and verisimilitude to exactly specify the situation with respect to which the required judgment is to be made. Such details can often be fixed across all scenarios under consideration.

Consider the evaluation of alternative naval aviation plans, and in particular, suppose we desire to compare two particular types of fighter planes, a VF1 and a VF2. A possible method would be to evaluate a particular sized fleet of ships consisting of specific numbers of carriers, destroyers, frigates, and the like. The carriers have specific numbers of squadrons of different types of aircraft systems. The only variable in the composition of the U.S. contingent to be evaluated is the identity of the fighter class, VF1 or VF2.

This fleet combination is evaluated in each of a number of operational settings involving different levels of enemy threats in different world areas. The utilities resulting from the outcomes of deploying the different fighters in the different settings could be compared and a deployment decision reached. Such an evaluation would require that the variable scenario details formed a valid representation of the future

world, and that these details also captured the relevant utility differences. However, it should also be the case that the relative utility differences should not change if the composition of the fixed U.S. fleet changed. That is, although the outcomes and associated utilities could change as a function of U.S. fleet composition, the relative utility differences would not change. If the latter were the case, the fixed set of scenario details, the U.S. fleet composition other than the VF class composition, is not relevant to utility differences and the results could be generalized to other fleet compositions. Even a small amount of military expertise will convince the reader that such assumptions must be carefully checked, for they are likely to hold for only limited parts of S, and the nature of the limitation must be specifiable.

Further, note that for purposes of expected utility calculation, scenario probabilities must be assessed. By restricting the evaluation to a particular U.S. fleet composition, a subspace of S is being considered, and the probabilities of different operational settings involving different levels of enemy threats in different world areas is conditional on the subspace of S under consideration. The relative values of these conditional probabilities should not change markedly as a function of the specification of the subspace of S. (Note, of course, that the absolute probabilities will change, for example, if the fleet has no carriers.)

There are two important purposes of fixing details across scenarios. One, already explained, is for clarity and verisimulitude, to level plausibility to a scenario and thus enhance probability estimation. A second is to reduce the eventual number of scenarios to be evaluated. If the number of variable details for scenarios is large, the

number of potential scenarios increases dramatically. In
the naval aviation example, for instance, the level of
hardware technology used combined with the identity of the
personnel manning the weapons was the most utility-relevant
dimension of the partitioning. But the probabilities of the
possible levels of this compound dimension were difficult to
assess; intuition was confounded. When presented with such
a question, an expert might well say, "It depends on what
part of the world you are talking about." Inclusion of the
part of the world being considered in the partitioning,
while not very utility-relevant, was worthwhile because it
made the probabilities of the cells of the partition of S
much easier to assess.

Even modest partitioning processes can often create
large numbers of scenarios, as the 120 scenarios of the
naval aviation example illustrate. It is easy to eliminate
preposterous scenarios. But what happens if after that
elimination too many are still left? This is essentially
identical with the problem of too many conditions in a full
factorial experimental design. So are its solutions. Random
sampling of conditions is an unattractive possibility.
Reduction of the number of dimensions of the design is
highly desirable, if feasible. Otherwise, the use of various
kinds of incomplete designs in which cells are deliberately
omitted in some orderly and systematic way seems like the
best solution. We have little new experience with this sort
of problem and, therefore, little to suggest beyond what is
found in any good textbook of experimental design. A major
point to keep in mind is that conditions identified by, say,
single-threading that are not relevant to discrimination in
utility among action options are prime candidates for elimi-
nation from the set of scenarios to be used. In this problem,
unlike many experimental design problems, we know ahead of

time which independent variable we really care about; it is
O. Other dimensions are helpful or relevant only if they
contribute to the discrimination among levels of O.

4.4 Evaluating Scenario Probabilities

In a very strict and useless sense, scenario probabilities
are all so near to zero that the task of assessing them is
futile. In the amount of detail usually added to a scenario
for plausibility and verisimilitude, any possible future is
preposterously unlikely, simply because there are so many of
them. (Consider the probability that your Social Security
number should have turned out to be exactly what it in fact
is.) But every event probability is defined over a universe
of events, and the universe relevant for our purpose is not
the universe of all possible futures, but rather the universe
of scenarios we plan to use. We have carefully chosen the
scenarios included within that universe not to be prepos-
terous and so implicitly considered the universe of all
possible futures. To calculate expected utilities, we need
probabilities of the scenarios we actually intend to use;
these probabilities will, by definition, sum to 1.0.

One approach, as unattractive as it is simple, for sce-
nario sets defined by partitions over several variables,
would be to treat the variables as independent, assess a
probability distribution over each level of each variable
separately, and then multiply the resulting dimension proba-
bilities to get the needed cell probabilities. If not all
cells in the Cartesian product are to be used, it is a
simple matter to renormalize so that the probabilities of
those that are to be used will sum to 1.0. The reason we
regard this approach as unattractive is simply that it
ignores probabilistic dependence between levels on different

41

dimensions. In the naval aviation example, the probability
of a crisis in the Mediterranean in which sophisticated
Soviet weapons are manned by unsophisticated non-Soviet
personnel is not the product of the probabilities of these
three levels of their respective dimensions taken separately;
it seems unlikely that sophisticated Soviet weapons will, in
fact, be manned by unsophisticated non-Soviet personnel.
(That is the reason why sophistication of weapons and of
personnel were combined into a single dimension in the
actual evaluation.)

There are two obvious ways out of the dilemma. One is
simply to consider each scenario separately, and to judge
its probability by any of the various techniques now used
for such purposes. This procedure is intellectually straight-
forward, but presents several problems. One is that the
number of judgments required may be large, depending on how
many scenarios are to be used. A second is that the procedure
makes no use of whatever dimensional structure may have been
involved in selecting the scenarios in the first place.

The other option is to judge appropriate sets of condi-
tional probabilities. In the naval aviation example, for
instance, one might start by judging the probability of
crisis in each of the six areas into which the world has
been divided. Then one might judge the conditional probability
of each type of conflict for each specific area of the
world. Finally, one might judge the conditional probability
of each level of Soviet technology and manning for a given
type of conflict and area of the world. Appropriate multipli-
cation will then yield the 120 scenario probabilities.
These can be renormalized as appropriate by treating the
probabilities of eliminated scenarios as zero. Compared to
the previous procedure, this one has both liabilities and

assets. It requires even more judgments than does direct
assessment of each scenario probability. But in using the
dimensional structure of the scenario set, it probably eases
the experts' job of making judgments.

So long as the partitioning on each dimension is, in
fact, exhaustive of S, this procedure can work. But that
partitioning may often not be exhaustive in the expert's
head, even if it is formally stated so as to be exhaustive.
The reason is that the expert may add additional detail to
the scenario in order to help his intuition. Instead of
thinking about crisis in the Mediterranean, for instance, he
may think of crisis in Lebanon. While such additional
specificity may aid intuition, it may also create intellectual
traps, in one or both of two ways. One occurs if intuition-
aiding detail is added to one level of a partition, but not
to other levels of that same partition. The other, more
subtle trap arises if intuition-aiding detail, added to all
levels of a partition, affects the probabilities of the
elements of the partition. Thus, an expert might regard a
crisis in Lebanon as a rather likely example of a crisis in
the Mediterranean, and a crisis in the Philippines as a less
likely but still appropriate example of a crisis in the West
Pacific. Use of these examples to aid intuition could make
judgments concerning the probability of a crisis in the
Mediterranean unduly high as compared with judgments of the
corresponding probability in the West Pacific.

If the expert uses intuition-aiding detail, he must be
alert to either sort of trap. The probability associated
with a scenario s_i should be the probability of the entire
region of S that that scenario represents, not just of some
portion of that region. Note that if the judge uses direct

43

assessment of scenario probabilities, he will probably find it much more difficult to guard himself against these two simply because he is not considering any formal partitioning of S.

When an incomplete subset of the Cartesian product of partitioning variables is used to define the set of scenarios to be used, renormalizing can present problems similar to those resulting from intuition-aiding detail discussed above. The omitted cells may introduce a bias into the assessment. However, the choice of cells for omission will ordinarily have been done to minimize any such effects. The most common reason for omitting cells is that they are preposterous; cells omitted for that reason can have no significant effect on relative utilities of options. The reason for omitting non-preposterous cells is simply to reduce the magnitude of the judgmental task in the analysis. Whatever procedure is used for the purpose, it will presumably have the goal of the analysis in mind and so will have been selected to minimize the extent to which the cells left in misrepresent the utility structure of (S x O). It makes little difference if they misrepresent the probability structure of S, so long as that misrepresentation does not lead to misrepresentation of the utility structure on (S x O).

The following procedure illustrated by the Navy air system mix evaluation problem depends on successful control of the possible biases discussed.

Recall that attributes were area (six levels) conflict type (5 levels), and the technology/manning factor (4 levels) partitioning S into 120 sub-units. This number is obviously too large. The expert can reduce the magnitude of the

problem by trying to create examples of each and thus locating those areas of S with zero probability. A more reasonable procedure involves use of a decision analytic probability diagram such as that illustrated in Figure 4-6.

The expert has chosen Lebanon as an example of the branch of the probability diagram shown by the darkened arrows. Certain branches of the diagram are impossible and therefore have zero probability. This probability can be ascertained by direct questioning of the expert or by attempting to create scenarios to represent the branch. For those branches with finite probabilities, the expert can think of several examples, the number being proportional to the probability of the branch. A reasonable procedure, therefore, is to have the expert, paying strict attention to those general scenario attributes, area, conflict level, and technology/ personnel, create as many scenarios as come to mind.

The probability of a branch could be determined as follows.

o   The relative probabilities of the scenario examples that characterize a branch are established. These probabilities will sum to 1.0 for each branch.

o   The probability of a branch is estimated by considering the product of the respective conditional probabilities of the events that comprise the branch. For the branch illustrated by darkened arrows in Figure 4-6 this probability would be the product of the probability of a conflict in the Mediterranean, the conditional probability of a

unilateral military action given a conflict in the
Mediterranean, and the probability of sophisticated
Soviet weapons manned by non-Soviet personnel
given a unilateral military action in the Mediterranean.

o   If all scenario examples are used for the analysis,
     the probability of the entire branch is partitioned
     among the scenario examples for the branch according
     to the relative intra-branch conditional probabilities.

o   If only two (or three) scenario examples are to be
     used for each branch, the probability assigned to
     the branch is partitioned among the scenarios pro-
     portional to the ratio of the respective intra-
     branch probabilities.

Another but mistaken procedure for probability assessment
is to establish scenario probabilities by considering the
relative pairwise likelihood ratios for scenarios. For
example, in Figure 4-6, s_1, s_8, s_{13}, and s_{17} might be chosen
to represent four different branches. Pairwise odds could
be assessed for each scenario pair and these could be converted
to scenario probabilities to be assigned to the appropriate
branches represented by the respective scenarios. This
procedure would be clearly misleading, for s_1 could be one-
tenth as likely as s_8, whereas the branch represented by s_1
might be four times more likely than results from that
represented by s_8. The solution offered here is one method
of handling this problem. As indicated, whatever procedure
is adopted, the nature of the required probabilities must be
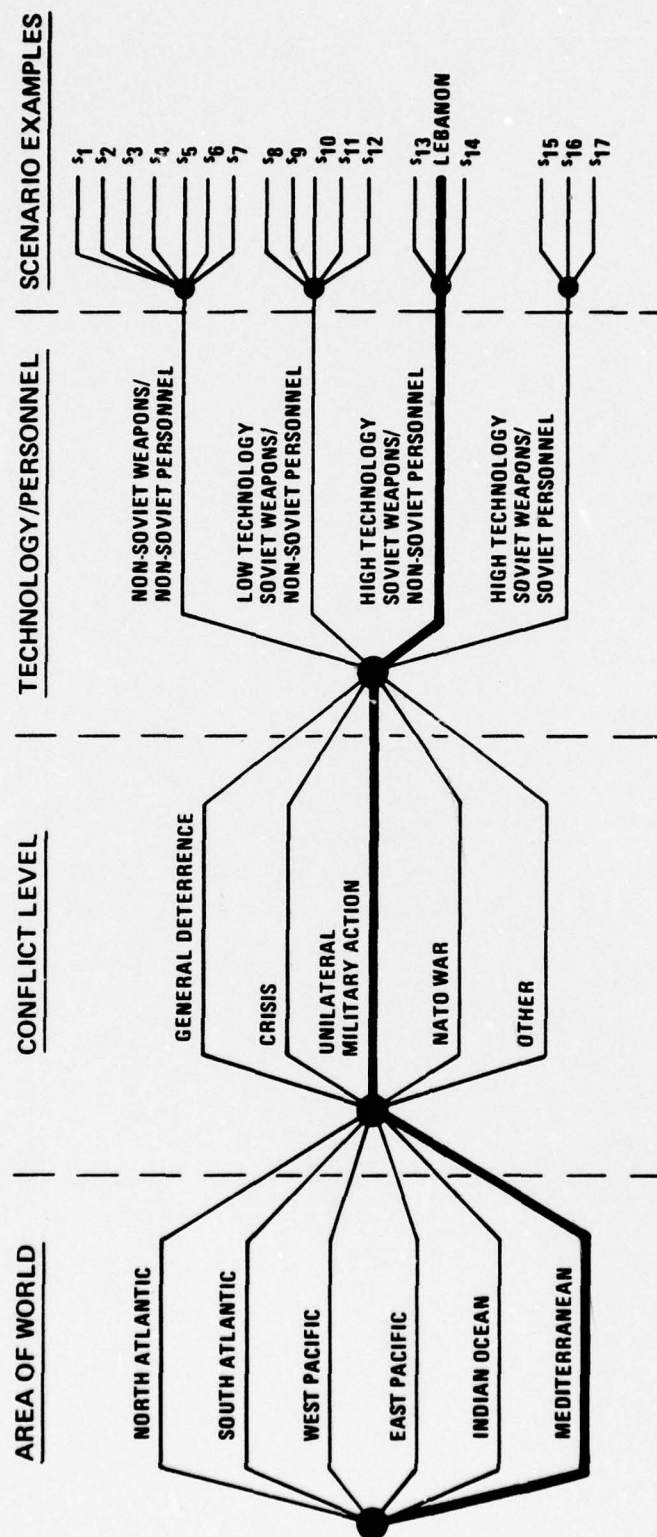kept in mind.

Figure 4-6

PROBABILITY DIAGRAM FOR THE NAVAL AVIATION EXAMPLE

4.5 Evaluation for Design: Tweaking in Option Space

Suppose that S has been partitioned into a representative
set of scenarios $\{S_i\}$, u_{ic} and p_{ic} have been assessed for
each s_i, and a need exists for a new or modified system.
The next step in the analysis is evaluation for design. O
must be examined to locate desirable systems, the attributes
of which can be used as a guide in specifying system character-
istics, inviting bids from suppliers, and the like.

Recall that the process of examining utility contours
in U as a function of (O x S) was likened to the turning of
dials in O, a process called tweaking. The purpose of tweaking
is to find as good a design as possible. Of course, since the
same design will seldom be optimal for all scenarios, the
expected utility criterion must be used. Tweaking is illustrated
in Figure 4-7, in which the utilities of o_c and o_k are examined
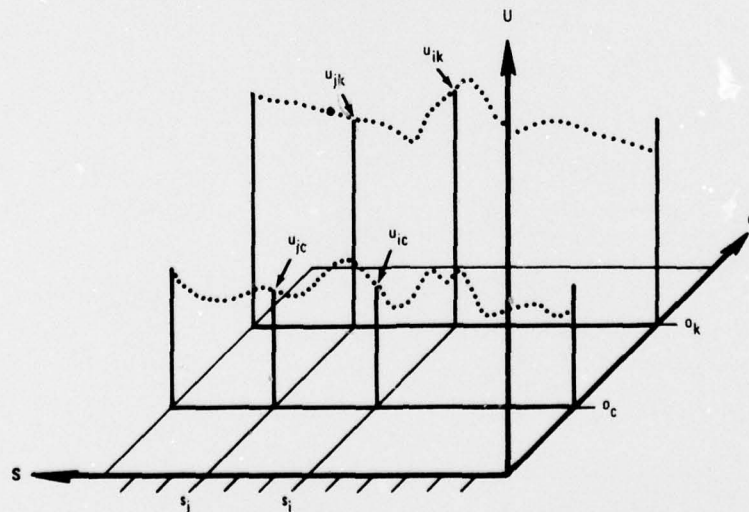across scenarios and are specifically compared for s_i and s_j.

Figure 4-7

THE UTILITY CONTOURS OF o_c AND o_k ACROSS S

48

For a particular s_i, we hope to find the utility contour of $(U \times s_i)$ by sampling O. The option with the highest utility for s_i is sought. Figure 4-8 illustrates such a hypothetical contour.
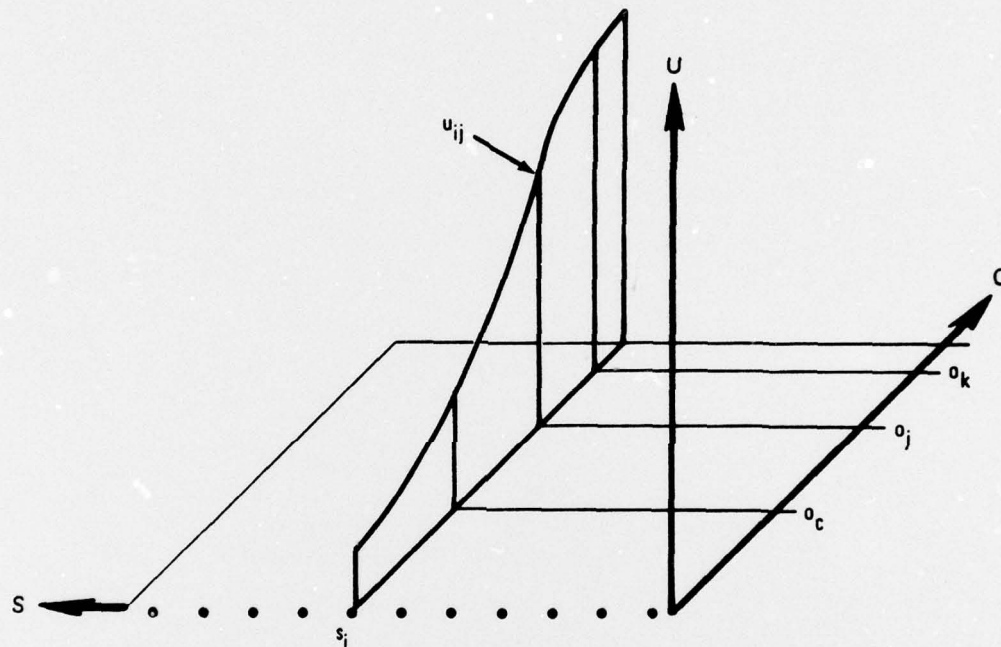
Figure 4-8

TWEAKING IN OPTION SPACE WITH s_i HELD CONSTANT

Another desirable property for the set of scenarios chosen to represent S arises here. For design purposes scenarios must be specific enough to allow tweaking to progress. Tweaking is a technical process that consists of examining projected performance and associated utility as a function of design parameters. If U as a function of the $\{s_i\}$ is insensitive to design changes, the scenarios must be altered, or the process must be modified.

49

A procedure is to examine P as a function of (O x S). Deficiencies of o_c in each s_i can be traced to performance characteristics, and optimal performance for each s_i can be established by examining (U x s_i x P). Such an analysis yields ranges of performance on the attributes of performance space along which system o_c is deficient. Design for s_i can then proceed by tweaking in O and observing the resultant effect in (O x s_i x P) by paying particular attention to the attributes of performance on which o_c was found to be deficient for s_i.

Given a specific s_i and specific performance require-ments, the potential number of combinations of option attri-butes that would yield the desired performance is very large. Since, in fact, the number of options in O is essen-tially infinite, each o_j representing a particular combination of attributes at whatever level of system attribute specificity has been chosen for analysis. Tweaking proceeds in O in much the same manner that S was partitioned--by attributes. Tweaking as a function of two attributes of option space is illustrated in Figures 4-9 and 4-10.

In Figure 4-9, as dial 2 (corresponding to option attribute 2) is tweaked for a fixed level of dial 1 (attribute 1), no change in u_{ij} occurs. Attribute 2 does not differentiate in terms of utility, and tweaking therefore progresses along attribute 1 with attribute 2 fixed at level 1. Utility does increase with increasing level of attribute 1. It appears that utility increases only as a function of attribute 1, but because the effects of dials will generally not be independent, the two points of Figure 4-9, the points associ-ated with levels 4 and 5 of attribute 2 with level 4 of attribute 1 fixed, are checked for differences from the utility of the pair of the points associated with level 4 of
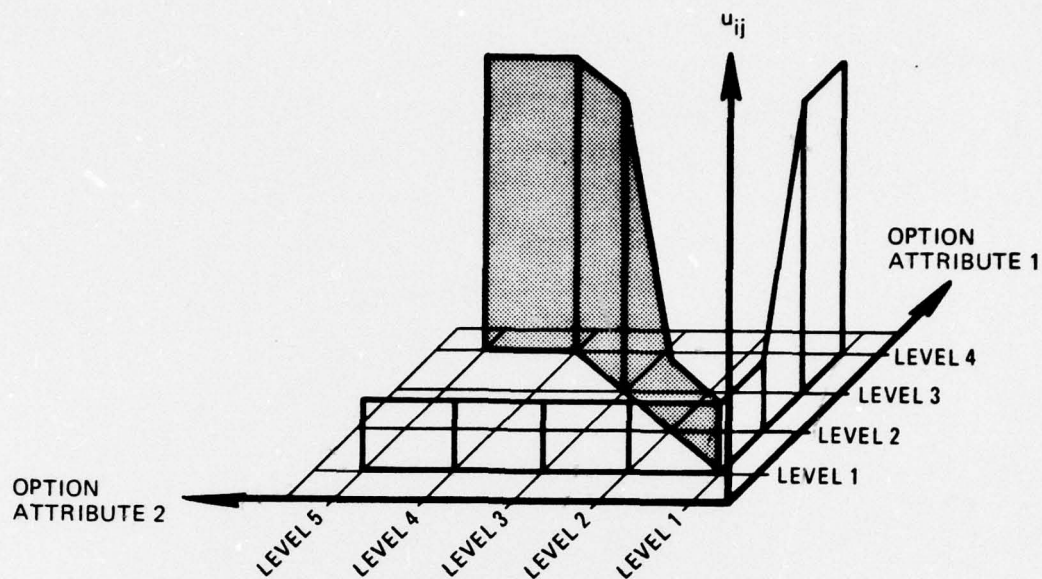
50

Figure 4-9

TWEAKING ON TWO ATTRIBUTES OF
OPTION SPACE IN SCENARIO s_i

attribute 1, level 1 of attribute 2. Since no difference
occurs, it may be reasonably assumed that the optimum for
these two attributes has been found. In Figure 4-10, turning
each dial with the other fixed at level 1 indicates that
utility is sensitive to both design parameters although it
is more sensitive to option attribute 2. Therefore, the
path illustrated by the arrows is followed to reach the
point of maximum utility.

An important question concerns the degree of specificity
of the dials chosen for tweaking in O. The degree of speci-
ficity of the dials depends on the degree of specificity of
the performance requirement. For example, the evaluation of

51
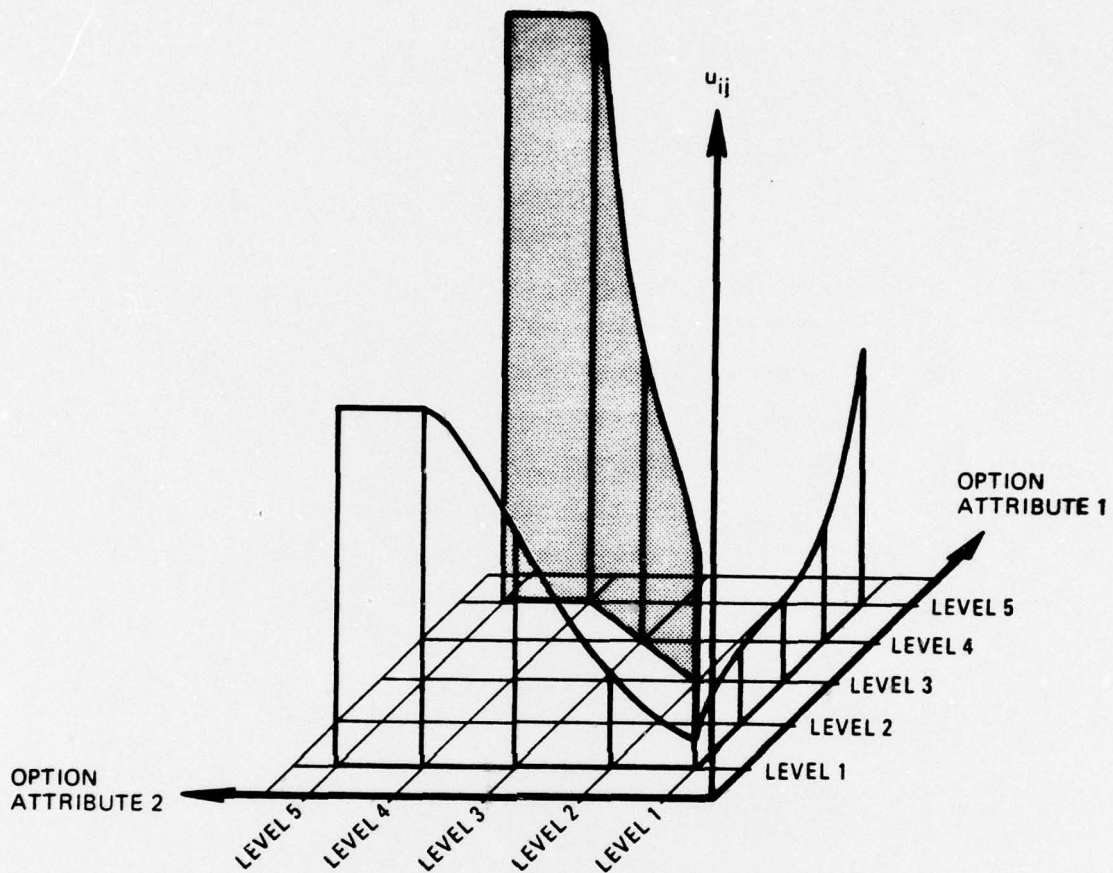
Figure 4-10

**TWEAKING ON TWO ATTRIBUTES OF
OPTION SPACE IN SCENARIO s_i**

the design of an EW system for defending a U.S. Navy ves-
sel may find that the current system is unable to defend
against a specific type of Soviet missile with a
particular radar. Appropriate defenses are either to
jam the seeker by sending a strong burst of radiated power
or to deceive the seeker. Design should therefore proceed
by turning either or both of the jamming and deception
dials. The deception dial is multi-attributed, for deception
can be accomplished in two ways. One is to create an

apparently larger target than the ship by using a decoy, a large chaff cloud consisting of metallic particles that are picked up by the seeker radar. The second is to return an electronic signal that makes the ship appear to be elsewhere. Tweaking with respect to deception can use either or both of these dials.

Suppose now that the scenario is modified by changing the nature of the threat slightly. The incoming missile is a heat seeker that looks for large sources of radiated heat. The design dials that would defeat the radar seeker will not work here. Thermal energy must be used either to jam or to deceive the missile. Thus, new dials in option space must be tweaked to cope with this threat.

Utility contours are thus examined over (O x S) to locate points of maximum utility. What kinds of problems can occur? The problem of local maxima is one example. A local maximum is a point o_j in (O x s_i) with utility greater than or equal to the utilities of the set of points near it, but not maximum for the entire space. This problem, sometimes difficult to handle, should be helped by reasonably extensive and systematic coverage of (O x s_i).

A second potential problem is that of sampling (O x s_i) by using s_i that are too broad for purposes of locating optima. Recall Figure 4-4 in which the utility of option o_c was examined as a function of two attributes of S. The utility contour, flat with respect to attribute 1, rises as a function of levels of attribute 2. Suppose, however, that attribute 2 was not isolated, that is, suppose that the nine scenarios were reduced to three by pooling the triads (s_1, s_2, s_3), (s_4, s_5, s_6), and (s_7, s_8, s_9). The utility contour for o_c as a function of these three more general scenarios

would be flat, and the sensitivity of design to scenario attribute 2 would be missed. This result suggests the general rule is that the path in S chosen for tweaking purposes should be narrow. Scenarios should be very specific. If the utility of designs turns out to be insensitive to certain attributes of S, these can be dropped, and scenarios can be broadened. Note, however, that such scenario attributes as the nature of the seeker in the previous example, if ignored, can lead to non-optimal designs.

The specificity of scenarios for design is, of course, dependent on the general nature of U as a function of (O x S) . If the situation is of a Kansas type, then sampling locates those few attributes to which U is sensitive. The utility contours as a function of other scenario details will be fairly flat, and any reasonable set of these other details can be chosen to yield a set of specific scenarios for purposes of estimating utility. The Kansas nature of the situation ensures that the resulting utilities for these scenarios are quite representative of the entire space.

When U as a function of (O x S) is of an Idaho type, U is very sensitive to scenario attributes, utility contours are narrow, more precision is required, and highly detailed scenarios are in order. Since the contours may and typically will include many local maxima, the number of scenarios necessary to capture utility variations is much larger than in the Kansas situation. That Idaho is a more serious problem than Kansas is simply a fact of life. If, in order to make the problem manageable, the set of scenarios chosen for evaluation for design is purposely reduced to a modest size by considering only certain types of threats and/or situations, those scenarios must be chosen with great care.

The result of this evaluation for design process is a
set of scenarios that may or may not be the same as those
for which the utility of o_c was evaluated to establish a
need. If the scenarios differ, this difference will be the
result of changes like added detail for design purposes,
reduction of the number of scenarios by choosing not to
redesign with respect to them, and the like.

With respect to option design, the results of evaluation
for design will be a general system description specifying
ranges of performance characteristics that are desirable,
specific required performances or capabilities in scenarios,
and a set of priorities on scenarios. The analysis, if
properly conducted, will allow statements about the required
capabilities in scenarios to be clear and, much more important,
consistent across scenarios. A system design is intentionally
somewhat abstract and specifies a set of constraints on the
proposed specific designs. Those constraints must be clear,
consistent, and easily translatable into detailed hardware
and procedures.

4.6 Evaluation for Choice

Once the procurement process has reached the stage of
choosing one or more designs for prototype development from
among those submitted, the major hurdles in the evaluation
for design and choice process have been negotiated. The
evaluator has a set of scenarios that represents the future
in which the system must operate and also should differentiate
among potential designs. Similarly, the utility of the
current system, o_c, has been evaluated in these scenarios.

Suppose that n proposals have been submitted specifying
n options $(o_1, o_2 \ldots o_n)$. Now the acts of deploying the

systems must be evaluated. The script must be generated for each o_j in each scenario. The evaluation of o_j in each scenario can proceed by direct judgment, simulation, or some combination of procedures. Again no procedure will be advocated, and comments on scenarios will be relevant to all procedures for estimating procedures.

Since the original set of representative scenarios contains those in which the need for new systems exists, these scenarios should be used for script generation. Two questions arise. First, how should this evaluation proceed? And second, should other scenarios also be used?

For o_c and $o_1...o_n$, and the set $\{s_i\}$ of representative scenarios, the evaluator must examine (U x S x O) by restricting the evaluation to the sub-spaces of O and S specified by the options and scenarios. As such, no new procedures are involved. The entire purpose now is to establish the expected utility differences between each of the proposed systems and o_c. A basic question, of course, is how well the proposed systems handle all threats in all scenarios. Each proposed system will, of course, have an expected utility in each scenario as well as an overall expected utility, but it is extremely unlikely that any system will be optimal in all scenarios. For each option, o_j, in s_i, the expected utility difference $p_i(u_{ij} - u_{ic})$ is calculated, and these differences are summed across scenarios to yield the expected utility difference or simply the expected utility of system o_j:

$$\text{E.U. } (o_j) = \sum_i P_i(u_{ij} - u_{ic})$$

Since it is often the case that the proposed systems have features that are designed to handle threats other than those represented by the $\{s_i\}$ chosen for evaluation, it can

be argued that evaluation of the $\{o_j\}$ using the $\{s_i\}$ may
fail to capture all aspects of the utility of $\{o_j\}$. Put
simply, the contractors may not always stick strictly to the
task of designing for the specific scenarios. Each design
will have certain features that the contractor thinks make
it the best of those proposed.

A second part of evaluation for choice can address this
issue. A fair way to do so is to design for each o_j a
scenario s_j, in which o_j performs as well as it possibly
can. The evaluation then proceeds by establishing the
relative probabilities of the resulting scenarios $\{s_j\}$
and the resultant expected utilities of each of the o_j by
using procedures already discussed. (The credibility of the
scenarios must, of course, be established prior to the
utility analysis.) Since the evaluator is no doubt tired of
the entire scenario problem at this point, and since the
evaluation of performance in the representative set of
scenarios is assurance of at least adequate future performance,
there is no reason not to allow each contractor to design
the scenario that is optimal for the system proposed by that
contractor. In fact, since no one else supposedly understands
the benefits of that system as well as that contractor, this
would seem to be an optimal procedure.

If the results of the two steps, evaluation in representa-
tive scenarios and evaluation in contractor-proposed scenarios,
agree on the best system, the problem is solved. If there
is some disagreement, the relative importances to be attached
to each set can be determined by the appropriate evaluating
authority, and choice can proceed.

57

5.0 CONCLUSION

This report has discussed the general topic of the use of scenarios in evaluation for the design and choice of complex systems. The discussion has been restricted to military systems, but in theory the problems and procedures discussed here are characteristic of any system evaluation problem.

The evaluation for design and choice proceeds from the original concept study in which a need for a system is established, through the evaluation for design in which system design requirements are established, to the final choice among alternative proposed systems. The need for a system must be established by first projecting a valid representation of the future in which the system will be deployed. Such projection involves the construction of a set of scenarios which has two properties. First, the scenarios jointly represent best guesses about possible futures. Second, the scenarios discriminate between alternative systems in terms of the utility of the deployment of such systems.

Once such a set of scenarios has been established, an obviously formidable task, the current system is projected into the scenarios for purposes of evaluating the resultant consequences of continued deployment of that system in the future. System shortfalls and resulting performance needs are established for scenarios. If a sufficient need for a new or modified system exists, alternative system designs are examined in scenarios to yield a set of system characteristics to be used in the design of proposed systems.

58

Proposed systems are then designed, and such systems are then evaluated against the representative set of scenarios to establish the expected utilities of the proposed systems for purposes of choosing one or more systems for prototype development.

There thus exist two aspects to evaluation for design and choice, evaluation of uncertainty and evaluation of utility, both of which must be accomplished to establish the desired evaluation metric, that of expected utility.

Too often the problem of addressing the uncertainty is finessed by stating that the task of validly representing the future is an impossible one. Thus, he who creates the scenarios can justify anything. Indeed, the sensitivity of expected utility to scenarios is obvious. It is difficult to conceive of a system for which one or two scenarios that justify the need for that system cannot be created. However, the set of scenarios proposed here is not such an arbitrary set. Rather it is a carefully constructed set which must, according to reasonable criteria, represent the future situations in which systems of the sort under consideration must operate. The designs are then created to satisfy scenario requirements and not vice versa. This point is crucial. Scenarios should be created prior to designs, and the characteristics of scenarios are design dependent only to the extent that the specificity of the scenarios must facilitate proper system design. The acknowledgment of the sensitivity of the evaluation problem to the nature of scenarios is, therefore, not a reason to argue against scenario use.

It should be obvious that the scenario topic is large and multi-faceted. This discussion has barely scratched the

surface of its problems. The procedures discussed here are essentially overviews or guidelines for looking at the scenario problem. Indeed, the purpose of this report is not to say what scenarios should be, but rather what they should do.

REFERENCES

Chinnis, James O.; Kelly, Clinton W., III; Minckler, Rex D.; and O'Connor, Michael F. Single Channel Ground and Airborne Radio System (SINCGARS) Evaluation Model. Technical Report DT/TR 75-2. McLean, Va.: Decisions and Designs, Inc., August, 1976.

Edwards, Ward. How to Use Multi-Attribute Utility Measurement for Social Decision Making. Technical Report 001597-1-T. McLean, Va.: Decisions and Designs, Inc., August, 1976.

Keeney, R.L. and Raiffa, H. Decisions with Multiple Objectives. New York: Wiley, 1976.

Wagner, Harvey M. Principles of Operations Research: With Applications to Managerial Decisions. Englewood Cliffs: Prentice Hall, 1969.